

Apr 20th, 12:00 AM - 12:00 AM

## Behind Derogatory Migrants' Terms for Venezuelan Migrants: Xenophobia and Sexism Identification with Twitter Data and NLP

Joseph Martínez  
*Old Dominion University*

Melissa Miller-Felton  
*Old Dominion University*

Jose Padilla  
*Old Dominion University*

Erika Frydenlund  
*Old Dominion University*

Follow this and additional works at: <https://digitalcommons.odu.edu/msvcapstone>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Data Science Commons](#), [Latin American Studies Commons](#), [Migration Studies Commons](#), and the [Social Media Commons](#)

---

### Recommended Citation

Martínez, Joseph; Miller-Felton, Melissa; Padilla, Jose; and Frydenlund, Erika, "Behind Derogatory Migrants' Terms for Venezuelan Migrants: Xenophobia and Sexism Identification with Twitter Data and NLP" (2023). *Modeling, Simulation and Visualization Student Capstone Conference*. 4.  
<https://digitalcommons.odu.edu/msvcapstone/2023/datascience/4>

This Paper is brought to you for free and open access by the Virginia Modeling, Analysis & Simulation Center at ODU Digital Commons. It has been accepted for inclusion in Modeling, Simulation and Visualization Student Capstone Conference by an authorized administrator of ODU Digital Commons. For more information, please contact [digitalcommons@odu.edu](mailto:digitalcommons@odu.edu).

# BEHIND DEROGATORY MIGRANTS' TERMS FOR VENEZUELAN MIGRANTS: XENOPHOBIA AND SEXISM IDENTIFICATION WITH TWITTER DATA AND NLP

Martínez, Joseph<sup>1</sup>, Miller-Felton, Melissa<sup>1</sup>, Padilla, Jose<sup>1</sup>, Frydenlund, Erika<sup>1</sup>

<sup>1</sup>Virginia Modeling, Analysis, and Simulation Center  
[Jmart130@odu.edu](mailto:Jmart130@odu.edu)

## ABSTRACT

The sudden arrival of many migrants can present new challenges for host communities and create negative attitudes that reflect that tension. In the case of Colombia, with the influx of over 2.5 million Venezuelan migrants, such tensions arose. Our research objective is to investigate how those sentiments arise in social media. We focused on monitoring derogatory terms for Venezuelans, specifically *veneco* and *veneca*. Using a dataset of 5.7 million tweets from Colombian users between 2015 and 2021, we determined the proportion of tweets containing those terms. We observed a high prevalence of xenophobic and defamatory language correlated with the terms, with the latter perpetuating a sexist portrayal of female Venezuelan migrants.

**Keywords:** Migration, xenophobia, natural language processing, social media.

## INTRODUCTION

Language is a powerful tool for communication can express hidden notions of power (Wanitzek 2002). Defamatory speech thrives on social media due the immediacy of responses and anonymity, which dehumanizes outgroups such as women and migrants (Haslam 2006). Xenophobic hate speech can spread rapidly online, creating a narrative that portrays migrants as villains rather than victims.

In this paper, we explore how historically ordinary words for Venezuelan migrants in Colombia, namely *veneco* and *veneca*, have evolved on Twitter to become markers of hate speech and xenophobic rhetoric. This study focuses on the usage of the terms' frequency and tone over time. Natural Language Processing (NLP) and a critical feminist lens are used to uncover how gender stereotypes contribute to hate speech on Twitter.

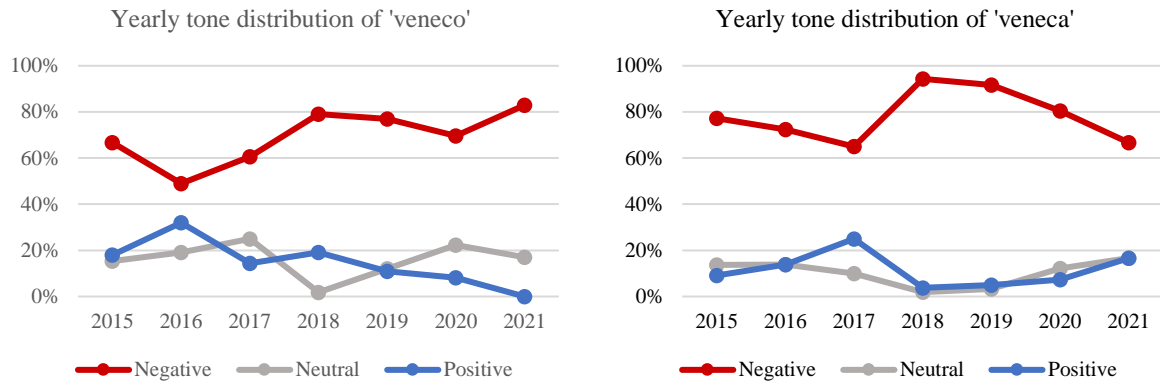
## MATERIALS AND METHODS

The study analyzed 5.7 million tweets from January 2015 to March 2021. A sample dataset of 4,684 tweets was labeled to determine relevance and tone, while another dataset of 3,996,268 tweets was geolocated using the Nominatim API. We developed a model to filter the complete dataset, and the labeled tone dataset was used to generate the frequency of the two terms across time. To accelerate processing, we trained a model to geolocate unmarked tweets, ultimately using 1,664,903 tweets in the study presented in this paper.

A sample of 4,684 tweets was labeled for relevance and tone classification. The tone was labeled along a Likert scale from highly negative to highly positive. NLP models were fine-tuned to filter the dataset automatically for both geolocation and relevance.

## RESULTS AND DISCUSSION

Notably, the proportion of tweets containing these terms is relatively higher in certain years, with a maximum of 25% and 8% for *veneco* and *veneca*, respectively, in 2015, followed by a decrease until 2018 and a rise again in 2021. However, relying solely on the tone proportions can be misleading, as it does not account for the context of each tweet. Figure 1 show the yearly tone distributions for both terms, which predominantly exhibit a negative tone. The negative tone has the highest peak in 2018, coinciding with a high increase in Venezuelan migrants entering Colombia.



**Figure 1. Yearly distributions of the tone of veneco and veneca, respectively.**

The term *veneco* was negatively impacted by events like soccer matches, political incidents, and comments by politicians. Deportations of Colombians by Venezuela in 2015 resulted in rejection and hatred comments. Meanwhile, *veneca* was frequently used to depict Venezuelan migrant women negatively, perpetuating harmful stereotypes, and contributing to their marginalization and stigmatization. Additionally, The use of *veneco* shifted in positive tone from 18% in 2015 to 0% in 2021, while negative tone has increased from 67% to 83% in the same period. This does not imply that the term has evolved from a benign term to hate speech based on assumptions about migrant behavior. Rather, negative perception has always been present, but now it is more prevalent and with less positive proportion.

## CONCLUSION

This study analyzed tweets using the terms *veneco* and *veneca* to understand public attitudes towards Venezuelan migrants in Colombia. The analysis of 1,664,903 relevant tweets showed that *veneco* was more frequent than *veneca*. Both terms had a high negative sentiment, with *veneca* often being associated with harmful stereotypes about Venezuelan migrant women. The prevalence of negative perception of the term *veneco* increased over time with a decrease in positive sentiment. Suggesting that the terms and the perception of the Venezuelan migrants have been always negative, with tendency to increase.

This study has limitations in that it only analyzed tweets with specific terms and relied on a small sample for tone analysis. Further research is needed to understand the impact of negative attitudes towards Venezuelan migrants on them. Overall, there is a need to create a more inclusive and supportive environment for Venezuelan migrants in Colombia.

## ACKNOWLEDGMENT

This research is funded by grant number N000141912624 by the Office of Naval Research through the Minerva Research Initiative; none of the views reported in the study are those of the funding organization. This research is funded by grant number 22RT0286 by the Air Force Office of Scientific Research through the Minerva Research Initiative; none of the views reported in the study are those of the funding organization.

## REFERENCES

- Haslam, Nick. 2006. 'Dehumanization: An Integrative Review', *Personality and Social Psychology Review*, 10: 252-64.
- Wanitzek, Ulrike. 2002. 'The Power of Language in the Discourse on Women's Rights: Some Examples from Tanzania', *Africa Today*, 49: 19 - 3.